Joining the dots: An automated method for constructing food webs from compendia of published interactions

Clare Gray [a,b], David H. Figueroa [c], Lawrence N. Hudson [d], Athen Ma [e], Dan Perkins [b], Guy Woodward [b,*]

[a] Queen Mary University of London, School of Biological and Chemical Sciences, London E1 4NS, UK
[b] Department of Life Sciences, Imperial College London, Silwood Park, Buckhurst Road, Ascot, Berkshire SL5 7PY, UK
[c] Universidad Católica de Temuco, Escuela de Ciencias Ambientales, Montt 56, Temuco Chile
[d] Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK
[e] Queen Mary University of London, School of Electronic Engineering and Computer Science, London E1 4NS, UK

ARTICLE INFO

ABSTRACT

Food webs are important tools for understanding how complex natural communities are structured and how they respond to environmental change. However their full potential has yet to be realised because of the huge amount of resources required to construct them *de novo*. Consequently, the current catalogue of networks that are suitable for rigorous and comparative analyses and theoretical development still suffers from a lack of standardisation and replication.

Here, we present a novel R function, `WebBuilder`, which automates the construction of food webs from taxonomic lists, and a dataset of trophic interactions. This function works by matching species against those within a dataset of trophic interactions, and 'filling in' missing trophic interactions based on these matches. We also present a dataset of over 20,000 freshwater trophic interactions, and use this and four well-characterised freshwater food webs to test the method.

The `WebBuilder` function facilitates the generation of food webs of comparable quality to the most detailed published food webs, but at a fraction of the research effort or cost. Furthermore, it matched and often outperformed a selection of predictive models, which are currently among the best, in terms of capturing key properties of empirical food webs. The method is simple to use, systematic and, perhaps most importantly, reproducible, which will facilitate (re-) analysis and data sharing. Although developed and tested on a sample of freshwater food webs, this method could easily be extended to cover other types of ecological interactions (such as mutualistic interactions).

1. Introduction

Characterising food webs (networks representing trophic interactions between species) and other ecological networks (networks which represent any type of ecological interaction, such as pollination) can help us understand and, ultimately predict multispecies systems' responses to changes in environmental conditions (Thompson et al., 2012; Tylianakis et al., 2010). Food webs can reveal subtle but important changes in the biotic interactions that underpin ecosystem functioning, stability, and resilience to perturbations — higher-level phenomena that cannot be inferred from studying the nodes (i.e., species or populations) alone (Gray et al., 2014; Thompson et al., 2012).

Despite the many advantages of a network-based approach to ecology, significant challenges need to be overcome, particularly in terms of gathering interaction data. Interactions occur between individuals and data are often collected at this level: for example, via collection, rearing and identification of every leaf miner, and subsequent leaf miner parasitoid along a transect to build herbivore–parasitoid networks (Macfadyen et al., 2011; as in Memmott et al., 2000), or through dissecting and identifying consumer gut-contents via microscopy (as in Layer et al., 2013). Such laborious methods require substantial investment of time and resources, and it can take many thousands of lab hours to characterise just one food web, which even then may still be undersampled for links between its rarer members (see Table 1; e.g. Woodward et al., 2005; Olito and Fox, 2014). Many hundreds or thousands of individuals of each species are often needed to fully characterise the full set of feeding links within a food web (e.g. Ings et al., 2009), which is rarely practical given the financial and time restraints of research funding. In addition, such comprehensive sampling is often destructive and can impose undesirable disturbance on study systems. Consequently, empirical food webs are often incompletely described and constructed from relatively small sample sizes (e.g. Kaiser-Bunbury et al., 2010; Layer et al., 2013). This limits the conclusions that can be drawn and the number of

* Corresponding author.
E-mail address: guy.woodward@imperial.ac.uk (G. Woodward).

Table 1
Methods for constructing food webs, with their advantages and disadvantages.

Method	Advantages	Disadvantages	Examples
Observation of evidence of interaction (e.g. feeding trials or gut contents analysis)	High confidence in links produced.	Very slow and labour-intensive. Rare interactions are often missed. Interaction type is biased by the method employed, e.g. the prey of suctorial predators cannot be determined through gut contents analysis.	Woodward et al. (2005) Macfadyen et al. (2011) Henson et al. (2009) Ledger et al. (2012)
Extrapolating from previously published interactions (e.g. `WebBuilder` function)	Fair confidence in links produced. Rare interactions can be included. Interactions from multiple studies determined through different methods can be easily incorporated. Low effort and quick.	Reliant on the quality of the data contained within the reference dataset. Can only be used to construct 'cumulative' or 'summary' food webs, i.e. temporal or spatial changes in feeding behaviour cannot be incorporated.	Hall and Raffaelli (1991) Goldwasser et al. (1993) Havens (1993) Piechnik et al. (2008) Pocock et al. (2012a); Layer et al. (2013) Cohen and Mulder (2014) Strong and Leroux (2014)
Predictive models	Ecological rules and theory can be incorporated. Low effort and quick.	Require prior knowledge of the structure of the food web in order to optimize parameter values. Many perform poorly at predicting individual interactions, even when food web structure is predicted well.	Cohen et al. (1985) Williams and Martinez (2000) Petchey et al. (2008) Allesina and Pascual (2009) Allesina (2011) Olito and Fox (2014)

comparable food webs that are available both across and within studies (Bascompte et al., 2003; e.g. Briand, 1983; Olesen et al., 2007) although exceptions to this exist (Bascompte et al., 2003; Cohen and Mulder, 2014). Most studies still have patchy and differing levels of sampling effort and taxonomic resolution, making meta-analyses difficult or even inappropriate: the ability to construct large numbers of realistic, comparable food webs across multiple systems would, therefore, help realise the true potential of network approaches (Gray et al., 2014).

Ecological networks are often constructed by incorporating species interactions from the published literature (Table 1) and many food webs are constructed entirely in this manner (e.g. Cohen and Mulder, 2014; Goldwasser et al., 1993; Havens, 1993; Strong and Leroux, 2014), while other food webs contain a blend of observational and extrapolated data (Layer et al., 2013; Pocock et al., 2012b). By filling in 'missing' trophic interactions to a given species list, the implicit assumption is made that, if a given pair of species have been observed to interact at one site, they will interact in the same way at other sites where they co-occur (at least in terms of a feeding link between the species being realised, or not). Food webs built through this method are often referred to as 'summary' or 'cumulative' food webs as they represent all potential interactions (of a particular type, for instance trophic interactions within a food web) between species of a particular community, rather than a snapshot in time. As such, food webs built through this method are unsuitable for detecting changes in species feeding behaviour across sites or over time, but are highly effective for detecting broad macro-ecological trends such as changes in food web structure across environmental gradients (e.g. Layer et al., 2010a,b; Mulder and Elser, 2009; Piechnik et al., 2008).

This approach can be taken further, by assigning interactions of species on the basis of taxonomic similarity: i.e., species within the same genus are assumed to have identical links if a link has been established through direct observation for at least one congener (e.g. Goldwasser et al., 1993; Layer et al., 2010a,b). This process is often used when constructing summary food webs for species the interactions of which have not been fully characterised (e.g., as revealed from yield-effort curves) to minimise potential biases arising from under-sampling, i.e. including *only* observed links would otherwise significantly underestimate food web complexity, especially among the rarer and/or more obscure taxa (Woodward et al., 2010). Recent work (Eklof et al., 2012) has provided justification for this approach, by highlighting the strong influence that taxonomy has in determining the structure of food webs. Thus, given the prevalence of undersampling in even relatively well-described food webs, dietary data extrapolated from the literature and generalised taxonomically can potentially produce far more complete and realistic

summary food webs than those that rely solely on observations made in a particular locale.

Despite the prevalence of these methods for constructing summary food webs in the literature (e.g. Cohen and Mulder, 2014; Goldwasser et al., 1993; Havens, 1993; Layer et al., 2010a,b; Pocock et al., 2012b; Strong and Leroux, 2014), there is still no standard method for inferring feeding interactions, resulting in inconsistencies among studies, even within the same ecosystem type. This is especially problematic because authors rarely state explicitly which links have been observed or extrapolated, or the source from which they have been drawn, or how closely the previously published interactions match those reported in their particular study, making replication impossible and preventing other researchers from scrutinising published interactions fully (but see Strong and Leroux, 2014).

Recent research has sought to develop predictive models of the structure of ecological networks (e.g. Eklof et al., 2012; Gravel et al., 2013; Olito and Fox, 2014; Rohr et al., 2010). Simple rules based on ecological theory have been used to model and predict the structure and topology of food webs, the most successful of which include deterministic models based on information on species' body sizes, for example the 'Difference', 'Ratio', and 'Difference/Ratio' models (Allesina, 2011) and the Allometric Diet Breadth Model (ADBM; Petchey et al., 2008) which incorporates allometric scaling and optimal foraging parameters. Whilst these models have been developed primarily to advance ecological theory, they provide a possible tool through which food webs could be built *de novo* in order to address questions about network structure across environmental gradients or scales. However, to achieve their best performance (proportion of correctly predicted links) these models require some prior knowledge about the number of links in the network. For instance, for the models mentioned above a researcher is required to go through a parameter optimisation procedure, by fixing the number of links, values of constants and exponents can be derived, by maximizing the number of links correctly predicted. When constructing a network for the first time for a particular system, a research would be required to fix the number of links to an expected value which would bias the network structure towards that which the researcher expected to find.

Additionally, recent work (Olito and Fox, 2014) has highlighted that while predictive models might perform well at predicting metrics of network structure, they tend to perform poorly at predicting pairwise interactions (e.g. Sáyago et al., 2013; Vázquez et al., 2009; Verdú and Valiente-Banuet, 2011; Vizentin-Bugoni et al., 2014), so whilst they may predict network structural metrics well, they are doing so for the wrong reason as the underlying biological mechanisms have not been

fully incorporated into the predictive models (Petchey et al., 2011). To the best of our knowledge, the models used here have not, up until now, been used to predict network structure *de novo*, as this is not the scenario for which they were developed.

Given the limitations of constructing food webs from observation of interactions or predictive models, we need an automated, repeatable and reliable method of building local food webs that can be applied across studies and, ultimately, different ecosystem and network types. Here, we introduce a method, the `WebBuilder` function that assembles food webs by systematically assigning links for taxa based upon a given set of user-defined rules applied to a dataset of known trophic interactions. We provide an implementation of our method for the R statistical modelling language (R Core Team, 2013), building upon the methods and data structures provided by the Cheddar R package (Hudson et al., 2013). We tested the method on four highly resolved freshwater food webs which have had their interactions characterised through gut contents analysis, as these represent some of the most complete food webs described to date, as a test case for our proof-of-concept. Specifically, our key aims were to:

1. collate a dataset of trophic interactions in a standard format to act as an example system in which to test this method;
2. automate the process of constructing food webs from this reference dataset in a repeatable and reliable manner;
3. compare the performance of this method with the structure of food webs with 'known' interactions, i.e. those which have been built through observation of the interactions; and
4. compare the performance of this method with another way of predicting food web structure; the ADBM, Difference, Ratio and Difference/Ratio models.

2. Methods

2.1. Dataset of trophic interactions

We collated a dataset of 20,823 pairwise trophic interactions among species (or the next highest level of resolution available, usually genus), from 51 different data sources, most of which were primary literature (Table A.1, Appendix A). It contains trophic interactions between primarily UK freshwater species, including 203 producer taxa, 593 invertebrate taxa, 24 fish taxa, 10,348 producer-animal links, 9531 animal-animal links and 944 detritus-animal links. When the necessary data were not available in the original publications, we contacted the authors directly, where possible, to obtain the raw data. The taxonomy of every resource and consumer has been standardised through the Global Names Resolver (http://resolver.globalnames.biodinfo.org/) using the Global Biodiversity Information Facility dataset. For every resource-consumer link the taxonomy (species, genus, subfamily, family, order, class) of both is given, along with life-stage information, if relevant, and a literature reference for the source of the link. This dataset builds upon the collection assembled by Brose et al. (2005), and to the best of our knowledge, represents the largest standardised collection of trophic links for freshwater organisms. This dataset is available to download at https://sites.google.com/site/foodwebsdatabase/ (doi:10.5281/zenodo.13751) and is designed to be easily updated by the iterative addition of new data (details of how to submit new data to the dataset are given on the website), allowing its content to improve over time, in an analogous manner to molecular-based bioinformatics datasets. New data will be subjected to a quality assurance procedure prior to inclusion in the dataset. Specifically, all taxa will be parsed through Global Names Resolver (http://resolver.globalnames.biodinfo.org/) using the Global Biodiversity Information Facility dataset. Additionally the data will be eyeballed for irregularities. It is anticipated that this data will exists as an open access resource, and as such the community of researchers who access it will report any errors they find so they can be double checked and removed. New iterations of

the dataset can be produced, hosted on the webpage alongside the original, and assigned a new DOI, allowing researchers to cite exactly which version of the dataset they have used for their research, allowing analyses to be repeated using identical versions to those cited in a given study, if required in the future.

2.2. The `WebBuilder` function

The method of constructing ecological networks by extrapolating from previously published interactions is implemented in a new R function – `WebBuilder` (see Appendix B for code). The user is required to provide the following; firstly a list of taxa (i.e. nodes) in the community of interest (step I, Fig. 1), this data can be gathered from multiple sources and could be in the form of survey or biomonitoring data. Secondly, for each node, the minimum level of taxonomic generalisation (explained below; step II), and the taxonomic classification of each node (step III). Lastly a registry — a dataset of known trophic interactions, including taxonomic classification (step IV), as example of which is published here, but which can also be created by the user or obtained elsewhere. It is recommended that the user resolve the taxonomy of their taxa list and registry using the same procedure so as to ensure that taxa are matched correctly, if the user were using the registry provided here they would need to parse their taxa list through Global Names Resolver (http://resolver.globalnames.biodinfo.org/) using the Global Biodiversity Information Facility dataset. The function searches the registry for every possible combination of resource-consumer interactions (for N taxa there are N^2 possible trophic interactions) which match the provided taxa list given the specified level of taxonomic generalisation.

The minimum level of taxonomic generalisation determines the taxonomic rank at which matches are made, thus generalising the resources or consumers of the candidate node to the species, genus, subfamily, family, order etc. level, as specified in the input (step II, Fig. 1). For instance, a researcher might decide to ascribe the level of taxonomic generalisation of 'genus' to the mayfly *Baetis fuscatus*, allowing it to be matched with the more commonly studied species *Baetis rhodani* in the dataset, and take on the appropriate feeding interactions of that species, i.e. those which include taxa also present on the provided taxa list (see the first Scenario in Fig. 2). This level of taxonomic generalisation is selected based on knowledge of a candidate node's trophic interactions in relation to its sister taxa (i.e. if all members of the same taxonomic unit can be assumed to have the same trophic interactions or not), and this can be tailored depending on the resource/consumer status of the node. For example, consumers of the larvae of the non-biting midge subfamily Tanypodinae tend to be trophic generalists and would likely consume other larvae of the family Chironomidae, while it is not likely that the resources of Tanypodinae larvae (which are predominantly predatory) would be shared by all Chironomidae larvae (many of which are grazers or filter feeders). Hence it would not be appropriate to assign the trophic generalisation level 'family' to both resource and consumer interactions of Chironomidae. Instead a researcher might ascribe the 'resource method' for Tanypodinae as 'family', but the 'consumer method' as 'subfamily' (see the second Scenario in Fig. 2). The function output contains references to the original empirical links, the number of matches that were found and the taxonomic level at which those matches were found, so links can be additionally screened and scrutinised *post hoc*, and analysis can be repeated easily because the function output contains the necessary information. Example R code is supplied (Appendix C).

2.3. Comparing the `WebBuilder` function with empirical food webs

The `WebBuilder` function was validated on a collection of highly-resolved stream food webs which have had their trophic interactions characterised through direct observation; Broadstone Stream (Woodward et al., 2010), Afon Hirnant (Gilljam et al., 2011; Woodward et al., 2010), Tadnoll Brook (Edwards et al., 2009) and the summary food web for the
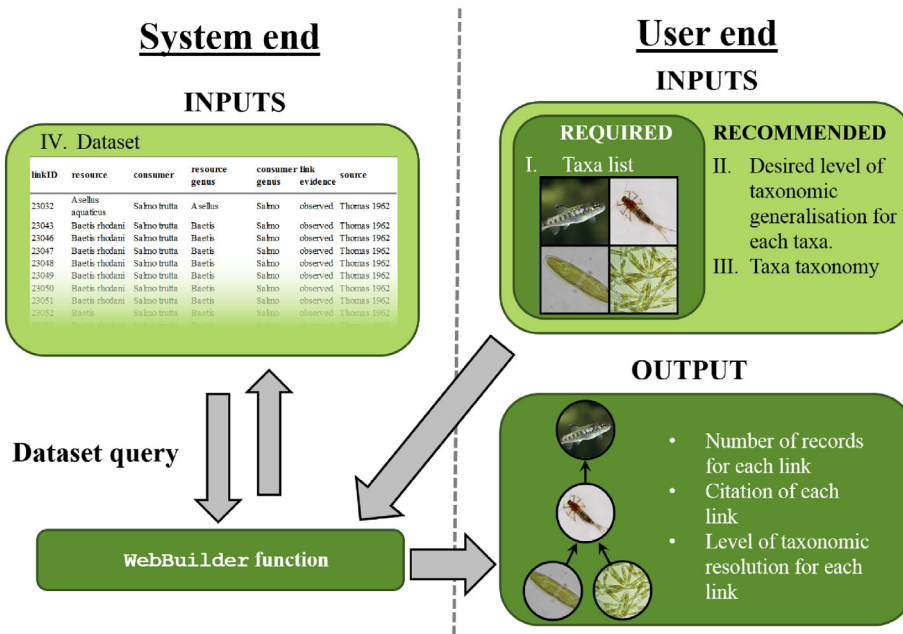
System end

INPUTS

IV. Dataset

linkID	resource	consumer	resource genus	consumer genus	link evidence	source
23032	Asellus aquaticus	Salmo trutta	Asellus	Salmo	observed	Thomas 1962
23043	Baetis rhodani	Salmo trutta	Baetis	Salmo	observed	Thomas 1962
23046	Baetis rhodani	Salmo trutta	Baetis	Salmo	observed	Thomas 1962
23047	Baetis rhodani	Salmo trutta	Baetis	Salmo	observed	Thomas 1962
23048	Baetis rhodani	Salmo trutta	Baetis	Salmo	observed	Thomas 1962
23049	Baetis rhodani	Salmo trutta	Baetis	Salmo	observed	Thomas 1962
23050	Baetis rhodani	Salmo trutta	Baetis	Salmo	observed	Thomas 1962
23051	Baetis rhodani	Salmo trutta	Baetis	Salmo	observed	Thomas 1962
23052	Baetis	Salmo trutta	Baetis	Salmo	observed	Thomas 1962

User end

INPUTS

REQUIRED

I. Taxa list

RECOMMENDED

II. Desired level of taxonomic generalisation for each taxa.

III. Taxa taxonomy

OUTPUT

- Number of records for each link
- Citation of each link
- Level of taxonomic resolution for each link

Dataset query

WebBuilder function

Fig. 1. A simplified workflow demonstrating the `WebBuilder` function. For a workable example see Appendix C.

INPUT

Taxa	Genus	Subfamily	Family	Resource method	Consumer method
Salmo trutta	Salmo		Salmonidae	exact	exact
Baetis fuscatus	Baetis		Baetidae	genus	genus
Tanypodinae		Tanypodinae	Chironomidae	family	subfamily
Navicula tripunctata	Navicula		Naviculaceae	genus	
Cocconeis placentula	Cocconeis		Cocconeidaceae	genus	

PROCESS - `WebBuilder` function & Dataset

Link ID	Resource	Consumer	Resource genus	Consumer genus	Resource subfamily	Consumer subfamily	Source
7081	Navicula lanceolata	Baetis scambus	Navicula	Baetis			Bert et al. 2008
7086	Cocconeis placentula	Baetis scambus	Cocconeis	Baetis			Bert et al. 2008
34136	Baetis	Trissopelopia	Baetis			Tanypodinae	Ernie et al. 2011
23047	Baetis rhodani	Salmo trutta	Baetis	Salmo			Bird et al. 1962
33415	Tanypodinae	Salmo trutta		Salmo	Tanypodinae		Ernie et al. 2011

OUTPUT

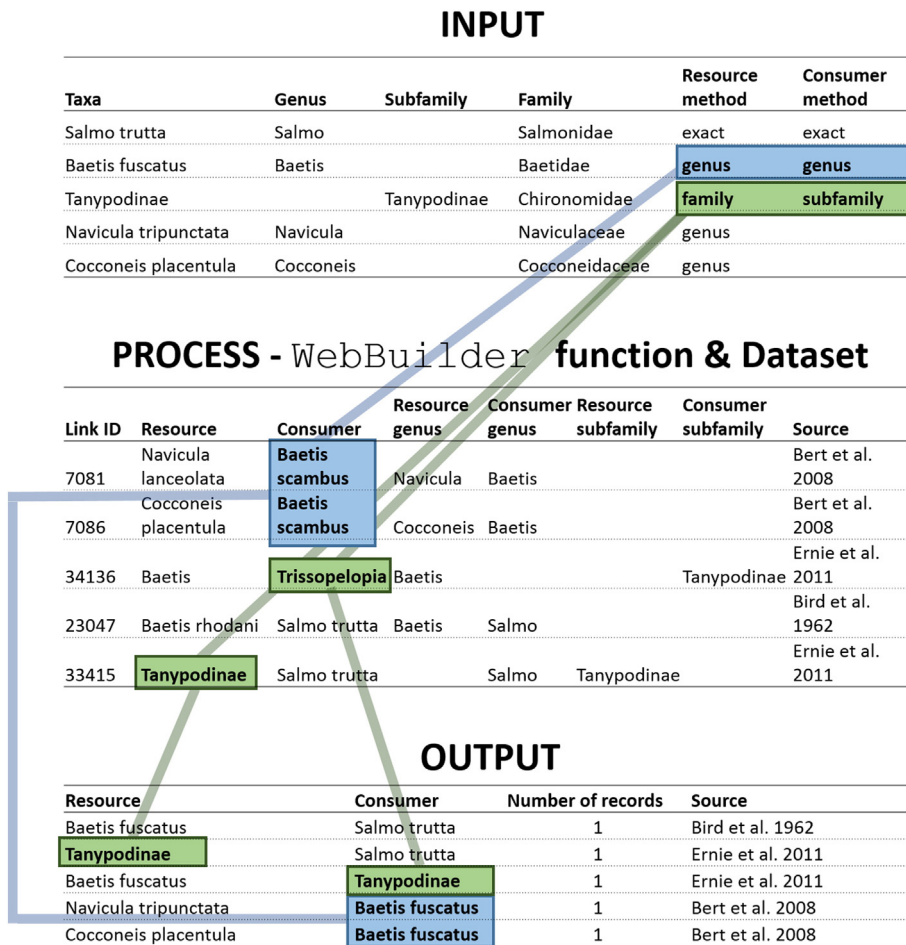| Resource	Consumer	Number of records	Source
Baetis fuscatus	Salmo trutta	1	Bird et al. 1962
Tanypodinae	Salmo trutta	1	Ernie et al. 2011
Baetis fuscatus	Tanypodinae	1	Ernie et al. 2011
Navicula tripunctata	Baetis fuscatus	1	Bert et al. 2008
Cocconeis placentula	Baetis fuscatus	1	Bert et al. 2008

Fig. 2. An example of inputs and outputs for the WebBuilder function. Two different scenarios are highlighted. Firstly in blue the taxa *B. fuscatus* is generalised to the genus level for both its consumer and resource links, this allows it to be matched with B. scambus in the registry and the Navicula tripunctata – *B. fuscatus*, and *Cocconeis placentula* – *B. fuscatus* links to be included in the output. Secondly in green, the taxa Tanypodinae are generalised to the family level for its resource links and subfamily level for its consumer links, allowing it to matched with all entries in the registry with the subfamily Tanypodinae and the Tanypodinae – *S. trutta*, and *B. fuscatus* – Tanypodinae links to be included in the output.

replicated four reference Mill Stream side-channels (Ledger et al., 2012; Woodward et al., 2012). The replicates for the Mill Stream data were pooled to aid comparison with the other food webs, which were all constructed as a single aggregate food web. The Broadstone and Afon Hirnant food webs contained only trophic interactions between macro-invertebrates, the Tadnoll food web contained interactions between macro-invertebrates and fishes and the Mill Stream data contain interactions between macroin-vertebrates, algae and detritus. When the `WebBuilder` function was used to generate the empirical food webs, in turn each respective local dataset was first removed from the global dataset, so each food web was generated in the absence of its own link information (to remove circularities).

The performance of the `WebBuilder` function was evaluated by calculating the True Skill Statistic (TSS; Allouche et al., 2006). This statistic was used as it can be digested into its component parts to provide information on the types of differences between the empirical and generated food webs, and builds upon the most commonly used metric which is simply the proportion of links correctly generated (e.g. Allesina, 2011; Petchey et al., 2008; Woodward et al., 2010). This statistic was chosen over likelihood based approaches because we were not interested so much in the efficiency of these predictive models, more the biological realism of the generated food webs (Petchey et al., 2011). The TSS is calculated from the following formula:

$$TSS = (ad - bc)/[(a + c)(b + d)]$$

where a is the number of links which were correctly generated by the function (the True Positives Rate; TPR), b the number of links generated by the function but not observed empirically, c the number of links not generated by the function but were observed empirically and d the number of links neither generated by the function nor observed empirically. TSS score values range from −1 to 1, where a score of −1 represents a generated food web that is the inverse of the empirically observed one (no observed empirical links are seen in the generated food web, and every non-link in the empirical food web is present in the generated food web), and 1 representing a generated food web having the exact same links as the empirically observed one.

Each empirical food web was generated using the level of taxonomic generalisation considered most appropriate (see Appendix D), this mostly consisted of exact and genus level matches although some family and order matches were used. To test how the generated food webs compared to their empirical counterparts, a series of network metrics were calculated; number of links (L), linkage density (L/S; where S is the number of nodes), connectance (C, where $C = L/S^2$), generality (the average number of resources per consumer), vulnerability (the average number of consumers per resource), and proportion of top, intermediate and basal nodes (with cannibalistic links removed). The difference between the generated and empirical network metric was tested with paired Wilcoxon signed rank tests.

To test how the quality of the generated food webs varied with dataset size, the dataset was randomly subsampled, in sequential steps of 5% from 5 to 100%, of the original dataset size, and then used to generate each food web. Each subsample size was repeated five times and each empirical food web was generated in the absence of its own food web data as above, to remove circularities. For each node within each network, the same level of taxonomic generalisation was used as above.

To test how the quality of the pairwise interactions generated by the `WebBuilder` function varied with the level of taxonomic generalisation, each food web was built using exact, genus, family or order taxonomic generalisation for all nodes. The degree (the number of links into or out of a particular node), generality, and vulnerability for every node in the generated food web was compared with that in the empirical network. The difference between the two for every node was recorded so that a positive score represented interactions 'missed' by the `WebBuilder` function, and a negative score represented 'extra' interactions not found in the empirical food web. The distribution of these scores gives an indication of how well the `WebBuilder` function predicted pairwise interactions

across the whole network: i.e., if, on average, it tended to 'miss' more interactions, or tended to pick up 'extra' interactions. To test if the mean was different from zero (indicating no difference in the quality of pairwise interactions between the generated and empirical food webs) a one sampled t-test was used.

2.4. Comparing the `WebBuilder` function to theoretical food web models

The performance of the `WebBuilder` function was compared with examples of some of the best-performing predictive models currently available: the 'Difference', 'Ratio' and 'Difference/Ratio' models (Allesina, 2011) as well as the Allometric Diet Breadth Model (ADBM; with "ratio" handling time, Petchey et al., 2008). The 'Difference', 'Ratio' and 'Difference/Ratio' models all generate food web links on the basis of body size, (either the difference between consumer body size and resource body size, the ratio between the two, or the difference multiplied by the ratio). The ADBM builds on this and incorporates allometries of body size and foraging behaviour of individual consumers to model food web structure (Allesina, 2011 for more detailed explanations; see Petchey et al., 2008). Detritus nodes were first excluded from the Mill Stream food web because these nodes had no body size or abundance data. For the 'Difference', 'Ratio' and 'Difference/Ratio' models two parameters required optimisation, a and b. For the ADBM we used parameter values for the mass to attack rate constant (a), resource mass to attack rate exponent (a_i) and consumer mass to attack rate exponent (a_j) from the literature (Rall et al., 2012) rather than through parameter optimisation as in Petchey et al. (2008), so as to simulate a situation for which the `WebBuilder` function was designed, where food webs are being generated for the first time with no prior knowledge of the system other than the species richness. For two parameters (mass to handling time constant, $h.a$; mass to handling time critical ratio, b) we were unable to find information in the literature with which to value these parameters, so went through the process of optimisation. For all models this was achieved by constructing food webs with a range of values for each parameter, and selecting those food webs which had a number of links that was within the range set by the `WebBuilder` function, i.e. if the `WebBuilder` function generated K links, and there were L empirical links and $K–L = t$ we selected all possible solutions within the range $L-t:L + t$, to make the comparison with the `WebBuilder` function fair. Note that for some food webs the difference between L and K was large, leading to large variation in the
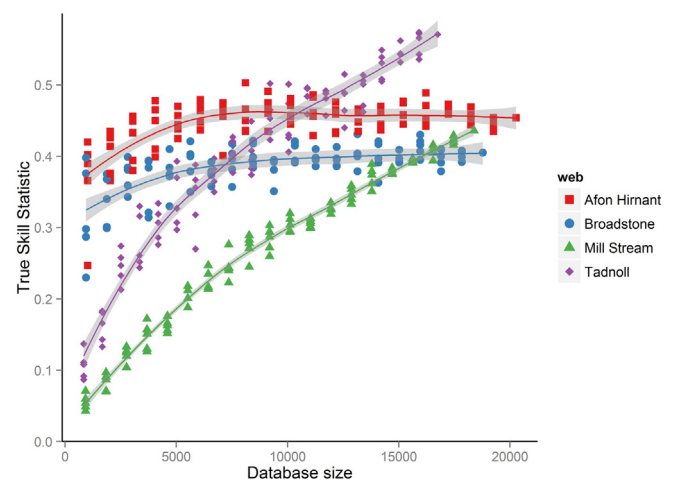

Fig. 3. The quality of the generated food web increases with the size of the dataset. Fluctuations in the TSS score are caused by changes in the component parts of the TSS, i.e. while the TPR may increase as the dataset size increases, other metrics such as the FPR might also increase, causing the total TSS to fall (see Methods). Lines are fitted using a LOESS smoother (Cleveland et al., 1992), grey shading indicates the 95% confidence intervals.

food web sizes generated by these models. Indeed for the Afon Hirnanlt food web this range fell below zero, and so the range was arbitrarily set to be the same proportional size as that of the Tadnoll food web, which had the next highest range. Parameter optimisation was conducted without using the connectance of the empirical food webs, hence although the same data have been used, results will vary from previous publications. Prior knowledge of the connectance of food webs would not be possible if a food web were being built *de novo*, so here we are using these models in a different way from their original application.

Fig. 4. Box plots showing the changes in generated trophic interactions as the level of taxonomic generalisation is varied. The difference in degree (top), generality (middle) and vulnerability (bottom) of individual nodes between the generated and empirical food web, thus the sample size reflects the number of nodes in the empirical food web. Positive values represent links which were 'missed' by the `WebBuilder` function, while negative values represent additional links not found empirically. Box plots are colour coded: Broadstone (B; blue), Afon Hirnant (A; red), Tadnoll (T; purple) and Mill Stream (M; green). Stars indicate if the mean is different from zero (one sampled t-test) and indicate if the generated trophic interactions are different from that of the empirical food web, $0.05 > = p > 0.01 = *$, $0.01 > = p > 0.001 = **$, $p < = 0.001 = ***$.

3. Results

3.1. Comparing the `WebBuilder` function with empirical food webs

When we constructed food webs from random subsets of the dataset, the quality (as measured by TSS scores) of the generated food webs improved as the number of records in the dataset increased, allowing more complete resource and consumer interactions to be ascribed to each taxa (Fig. 3). The strength of this relationship was food web specific, for instance Broadstone and Afon Hirnant did not continue to improve beyond a dataset size of about 25%. These food webs are relatively simplistic compared to Tadnoll and Millstream, and so the `WebBuilder` function reached its optimum performance when generating these food webs with a fraction of the total dataset. Tadnoll and Millstream did not reach their asymptotes suggesting that more data is needed to improve upon the quality of their generated food webs.

The level of taxonomic generalisation for each node was important for the quality of the generated food web; if the taxonomy of a given node list was generalised too far (typically beyond the family level) then the ascribed links became unrepresentative and the food web become over-connected resulting in an increased FPR and lower TSS score (Fig. A.1, Appendix A). At the scale of individual trophic interactions, the difference in degree, generality and vulnerability was generally positive when matching taxa exactly or at the genus level, and becomes progressively more negative as the taxonomic generalisation increased, indicating that the `WebBuilder` function was 'missing' links when matching nodes exactly or at the genus level, and including progressively more links the further the taxonomy was generalised (Fig. 4). For Afon Hirnant and Tadnoll there was no significant difference in the generality of consumers between the generated and empirical food webs when taxa were matched at the genus level, and there was no significant difference in vulnerability of resources for the Tadnoll food web when match at the genus level. This suggests that matching taxa at the genus level for these food webs produces the most 'accurate' pairwise interactions. For all other food webs and levels of taxonomic generalisations the generated links were different from that of the empirical food web (Fig. 4).

The occurrence of each food web's nodes in the dataset is given in Table 2. The coverage of these within the reduced dataset (food web data were removed from the dataset when used to generate the food web for that site) varied between 1497 (Broadstone) and 6704 occurrences (Mill Stream) (Table 2). Even at the family level some nodes from Broadstone, Afon Hirnant and Mill Stream were not represented in the dataset, meaning that those nodes needed to be generalised further still for the `WebBuilder` function to generate their links. These nodes tended to be rare taxa which were poorly represented in the dataset. The generated food webs (see Appendix S4 for the generated trophic links) had similar network metrics to the empirical food webs (Table 3), although the proportion of top nodes was consistently lower in the generated food webs, and the proportion of intermediate and basal nodes was consistently higher. All generated food web metrics were found to be similar to that of their empirical counterparts (paired
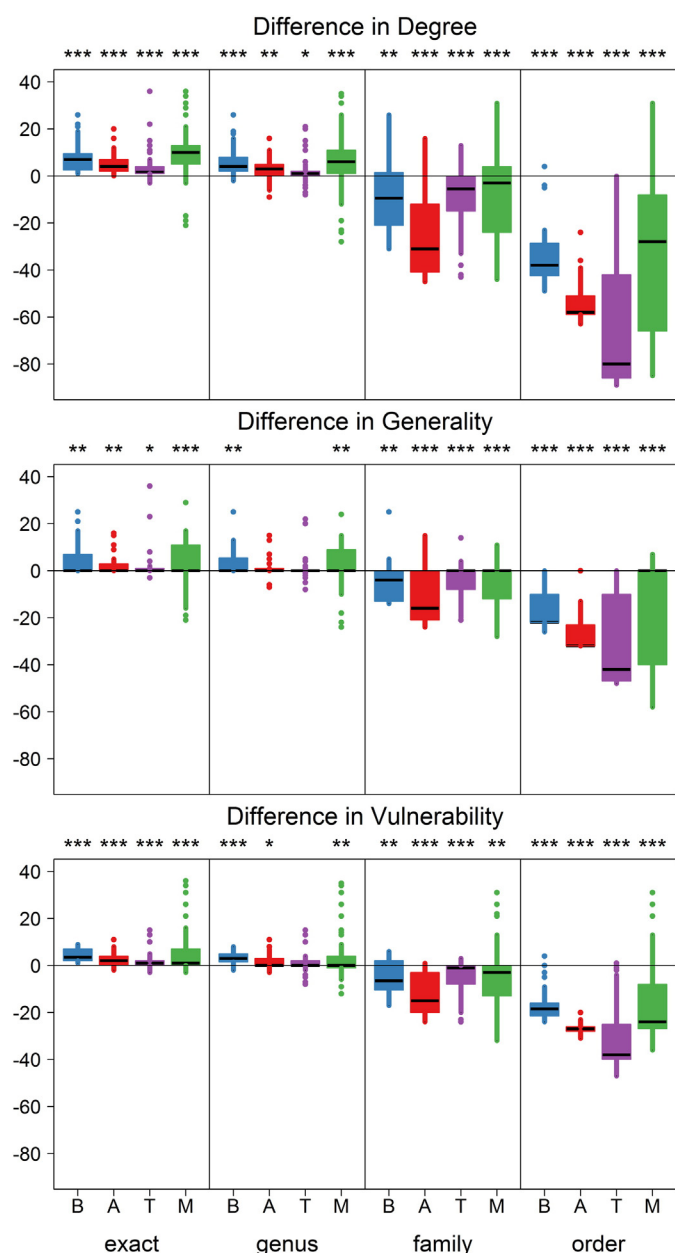
Table 2
The representation of the food web taxa within the full dataset and partial dataset (i.e., diet data gathered from a food web were excluded from the generation of its own inferred food web).

Food web	Number of appearances in dataset		Percentage of nodes appearing in partial dataset at each taxonomic level		
	Full data set	Partial data set	Exact	Genus	Family
Broadstone	2,196	1,497	81%	84%	94%
Afon Hirnant	2,945	2,266	72%	79%	92%
Tadnoll	12,405	4,314	84%	91%	100%
Mill Stream	11,545	6,704	87%	96%	97%

Table 3
The number of links (L), linkage density (L/S, where S = number of nodes), the connectance (C, where C = L/S²), generality, vulnerability, proportion of top, intermediate and basal species of the empirical and generated food webs. The performance of the WebBuilder function (relative to the original empirical food web) is summarised by the TSS statistic (which gives an overall measure of performance), and TPR (the proportion of links correctly generated). All food web metrics for the generated food webs were found to be similar to that of their empirical counterparts (paired Wilcoxon signed rank test, p = > 0.05).

Network		L	L/S	C	Generality	Vulnerability	Top	Intermediate	Basal	TSS	TPR
Broadstone	empirical	124	4.43	0.158	13	4.33	0.68	0.29	0.04	0.405	0.376
	generated	194	6.93	0.247	9.84	7.48	0.04	0.64	0.25		
Afon Hirnant	empirical	93	2.82	0.085	7.67	4	0.58	0.12	0.24	0.454	0.24
	generated	250	7.58	0.23	15.87	8.5	0	0.45	0.39		
Tadnoll	empirical	169	2.91	0.05	9.33	3.23	0.69	0.21	0.1	0.571	0.372
	generated	285	4.91	0.085	12.45	5.59	0.07	0.31	0.53		
Mill Stream	empirical	680	9.19	0.124	16.98	14.15	0.46	0.19	0.35	0.436	0.531
	generated	639	8.64	0.117	16.63	11.7	0.19	0.32	0.41		

Wilcoxon signed rank test, p = > 0.05). The TSS ranged from 0.405 (Broadstone) to 0.571 (Tadnoll Brook). Two food webs (Broadstone and Afon Hirnant) contained nodes that were found to have predatory links in the empirical food web but were not predicted to have any by the WebBuilder function, and vice versa many nodes distributed across all the food webs were predicted to have consumer links but were not found to have any empirically (Fig. 5).

3.2. Comparing the WebBuilder *function method to theoretical food web models*

The percentage of links correctly predicted (TPR) by the ADBM ranged from 12 to 43%, for the Difference model it was 0–46%, Ratio model it was 0–51% and the Difference/Ratio it was 0–54% (Fig. 6). All webs generated by the WebBuilder function had higher TPR and TSS
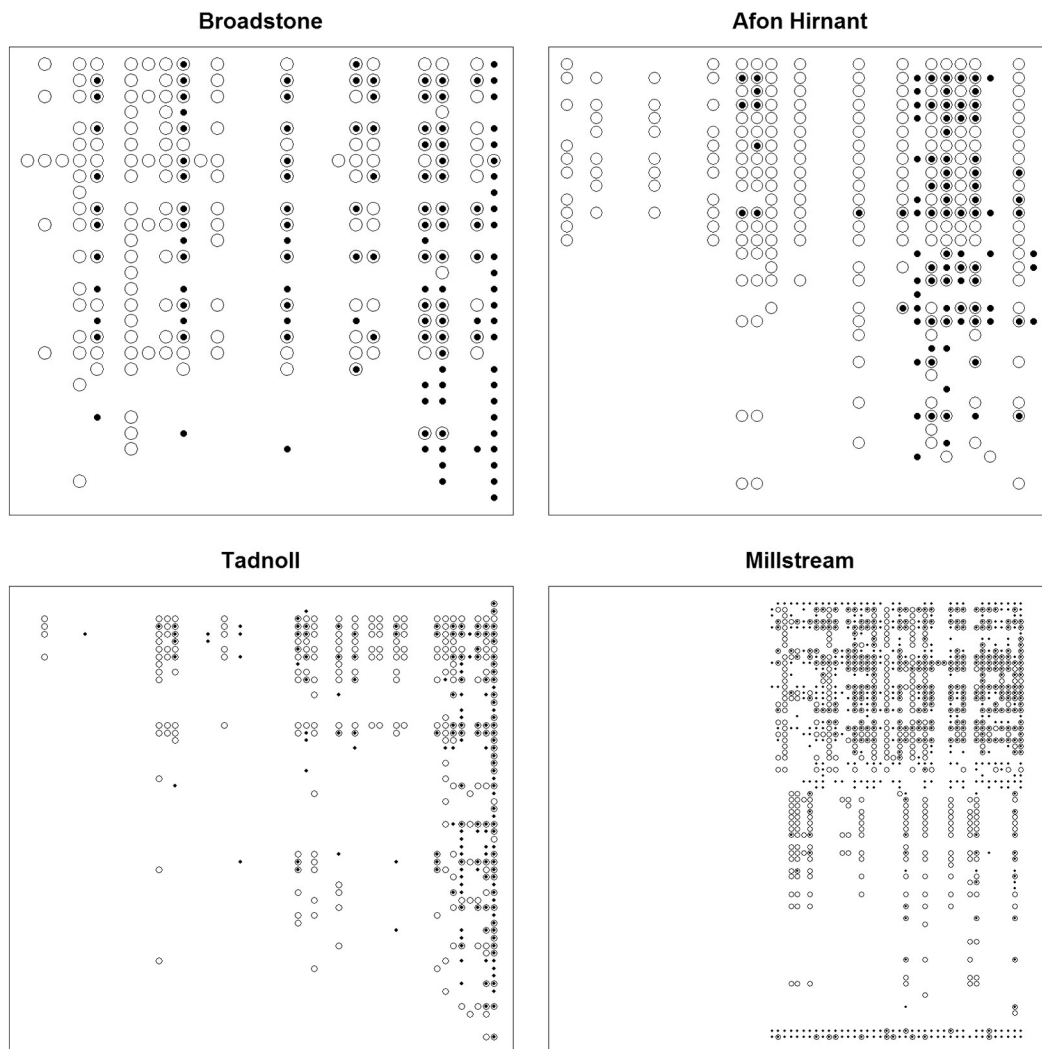

Fig. 5. Predation matrices for the empirical food webs compared to those generated by the AR method. Nodes are ordered by increasing body mass. A trophic link is represented by a point indicating that the taxon in that column consumes the taxon in that row. Links generated by the WebBuilder function are represented by empty circles, and those found empirically are represented by smaller, filled circles.
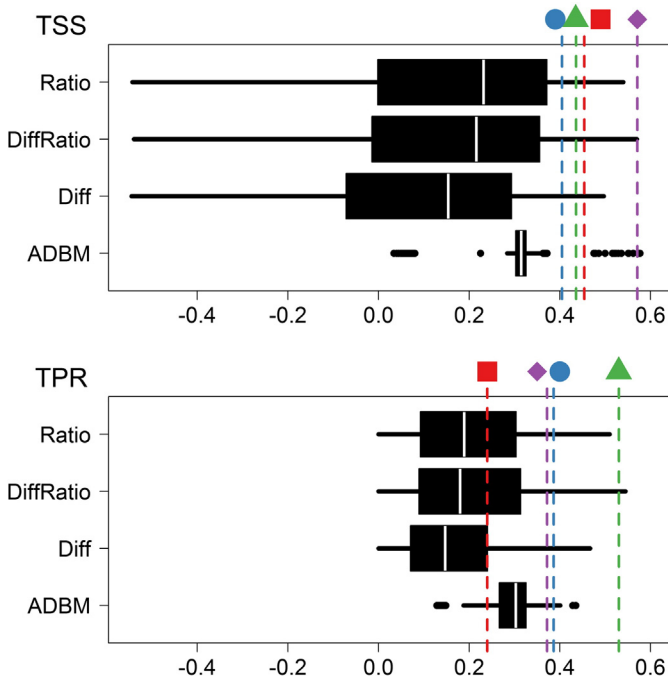
Fig. 6. Box plots showing the performance of the `WebBuilder` function compared to the ADBM, Difference, Ratio and Difference/Ratio models. The performance of the `WebBuilder` function is plotted as four vertical lines, one for each of the empirical food webs; Broadstone ●, Afon Hirnant ■, Tadnoll ◆ and Mill Stream ▲. The TSS score (top panel) gives an overall measure of the performance of the predictive method relative to empirical food webs, and varies between 1 (a generated food web that is exactly the same as the empirical food web) and −1 (a generated food web which is the exact inverse of the empirical food web). The TPR (True Positives Rate; bottom panel) is the proportion of generated food web links that were also found empirically, and varies between 0 (no links generated correctly) and 1 (all links generated correctly). A box plot of each set of values is given, indicating the range, quartile ranges and median of each set of values. For the `WebBuilder` function only the individual scores for the four food webs are shown, for all others there are too many generated scores to be shown individually; ADBM (n = 508), Difference (n = 32,025), Ratio (n = 43,638) and Difference/Ratio (n = 41,602).

scores than the median values for the Difference, Ratio and Difference/Ratio models (Fig. 6). In general the `WebBuilder` function had higher TPR and TSS scores than the ADBM, however the TPR score for Tadnoll and Afon Hirnant were similar to the median ADBM TPR score for that particular food web (as opposed to the overall median). Additionally the TSS score for Tadnoll generated by the `WebBuilder` function was similar to the median ADBM TSS score (Fig. A.2, Appendix A).

4. Discussion

4.1. Strengths and weaknesses of the `WebBuilder` function

Here we have demonstrated a systematic and reproducible method for building ecological networks from compilations of previously observed interactions. The `WebBuilder` function facilitates comparability across studies, re-analysis and data sharing. Although developed in the context of freshwater food webs, given its simplicity and generality the `WebBuilder` function could be easily applied to other systems, such as terrestrial food webs or even mutualistic networks. Plenty of other datasets already exist which could be exploited similarly to produce comparable, reproducible networks from marine and terrestrial systems (e.g., Barnes, 2008; Database of Insects and their Food Plants; http://www.brc.ac.uk/dbif/).

The `WebBuilder` function is an effective tool for constructing summary ecological networks for the first time. The overall performance (TSS) of the `WebBuilder` function exceeded that of the ADBM, Difference, Ratio and Difference/Ratio models. The proportion of correctly

predicted links (TPR) was similar to or exceeded that of the ADBM. The ADBM cannot predict links for nodes that have no body-size information – either because it is not known or because the concept is meaningless for the node, such as detrital resources. This problem does not apply to the `WebBuilder` function. The ADBM, Difference, Ratio and Difference/Ratio models have been used to generate the food webs presented here before, and have performed better than we have achieved here (Allesina, 2011; Petchey et al., 2008; Woodward et al., 2010), however to achieve this accuracy the generated food webs were constrained to have the same connectance as the empirical food webs, an approach not available when building a food web for the first time. Indeed there were instances here that the TPR and TSS of modelled food webs exceeded that of the `WebBuilder` function, but from the range of possible food webs generated by these models, it is impossible to select the most 'accurate' one without knowledge of the expected number of links. The `WebBuilder` function does not rely on prior knowledge of the food web, only on the correct identification of the nodes, thus reducing biases and restrictions.

It is perhaps unfair to compare the performance of the `WebBuilder` function to that of the ADBM, Difference, Ratio and Difference/Ratio models due to the inherent differences in the mechanisms through which they operate, and indeed it is not our intention for this exercise to be taken as a criticism of these alternative approaches. Rather, we have compared them here in order to place the WebBuilder method in the broader context of some of other more widely-used predictive methods currently available. Comparing our approach with the performance of the ADBM essentially represents a test of, and a means of improving, our understanding of the mechanistic theory behind these trophic interactions. Comparing the food webs produced by our approach with empirical food webs represents a test of the quality of the underlying data held within the dataset of trophic interactions. The `WebBuilder` function should be used as tool with which to construct large collections of food webs with which to test our understanding of food web structure across environmental gradients. The `WebBuilder` function is particularly suited to constructing food webs for data-poor systems, e.g., where there is no information available about the abundance or body size of nodes, with the only information being a list of species present. Clearly the ADBM or other predictive models would not be suited to these conditions, as they were never designed to work in this way. The WebBuilder function, however, would be able to generate reasonably realistic food webs if given a reference dataset of relevant trophic interactions. The `WebBuilder` function is adaptive, and can be improved upon over time; for instance, by increasing the size and coverage of the dataset of interactions. Hence it requires a substantial amount of data to perform well, unlike the predictive models analysed here. These types of methods can be viewed as complementary: a researcher might use both in conjunction in order to harness the advantages of both to better predict food web structure: indeed we envisage combining the `WebBuilder` function and other predictive approaches in parallel to build and understand food webs.

The four food webs presented here are among the most highly resolved and complete freshwater food webs published to date, yet the links are still under-sampled for many nodes (Woodward et al., 2010), due to methodological issues and logistical constraints on sampling effort. The `WebBuilder` function can help to overcome these issues. Firstly, it can take many hundreds of individuals to characterise a species diet (Ings et al., 2009) and thus the interactions between rare consumers and rare resources are often under-sampled. The `WebBuilder` function helps to overcome this as rare interactions need only be observed once in the dataset of previously published interactions in order to be incorporated into applicable food webs as they are constructed: i.e., potentially the "global diet" of a species is held within the dataset, and can be expanded in future data collections. Secondly, the method of observing interactions often limits the types of interactions which can be characterised; for instance, the prey of suctorial predators (which are especially common in terrestrial ecosystems) cannot be identified through traditional gut

contents analysis, but if characterised through other means (e.g. laboratory trials or molecular sequencing) they can be included in the dataset and incorporated into generated food webs. For instance, two suctorial predators in the Broadstone food web (*Platambus maculatus* and *Bezzia sp.*) did not have their guts analysed for predatory links in the original study (Woodward et al., 2005) and so had been previously excluded from the food web (e.g. Petchey et al., 2008; Woodward et al., 2010), these nodes would have been predicted to prey upon other species by the `WebBuilder` function. This is due to the `WebBuilder` function generalising the taxonomy of these nodes, and their subsequent appearances in the dataset, as other studies have characterised the diets of these taxa. Some links in the dataset of trophic interactions were known from just a single data source, e.g. *Cordulegaster boltonii* as a consumer of *Nemurella pictetii* is known only from the Broadstone food web. Therefore, when we excluded self-referential diet data, the `WebBuilder` function reconstruction of Broadstone did not predict a trophic link between *C. boltonii* and *N. pictetii*. We have not quantified how often this effect occurred. As with other open-source datasets, anomalies will be ironed out as the dataset is enriched with more observations as it grows, and its coverage will improve over time.

Besides constructing food webs *de novo*, the `WebBuilder` function could be used to standardise a collection of networks gathered from different sources prior to analysis. This would effectively standardise the sampling effort for included interactions (although not for species richness or taxonomic resolution) and would remove spatially or temporally explicit interactions (or lack thereof). If the analysis was concerned with the structure of summary food webs from different locations and habitat types then this might be an appropriate first step.

4.2. Future directions

The realism of links generated by the `WebBuilder` function could be addressed by assessing the number of times a particular interaction appears in the dataset, as well as the number of times an interaction could have occurred but did not (i.e. species found at the same site but not found to interact). If a particular interaction has been observed many times across many systems, it is probably reasonable to assume it also occurs at other sites where those species co-exist. However, if it has only been observed rarely, or at a site with very different characteristics than the one in question (for instance contrasting environmental conditions, or significantly different community assemblages) this assumption might not be so reasonable. As the size of the dataset continues to grow, evaluation of whether links are realised or not will improve over time.

The `WebBuilder` function is designed to construct summary food webs, and ignores potential behavioural shifts of species, hence it is unsuitable for constructing temporally or spatially explicit food webs. Additional data such as abundance information could be used to weight interactions, this would, for instance, reduce the weight of interactions between rare species reducing their influence on food web structure and increasing the realism of the resulting food web. There is an increasing body of literature detailing the importance of weak and strong interactions within networks (e.g. Berlow et al., 2004; De Ruiter et al., 1995; Vazquez et al., 2007) and a multitude of methods already exist for determining interaction strengths in food webs (see Berlow et al., 2004) some of which can be employed alongside the `WebBuilder` function. Thus, despite the 'coarse' nature of food webs built in this way there is much potential for their use in ecological research, and by combining them with models such as those presented here potential mismatches arising from behavioural shifts could be highlighted.

It would be straightforward for the underlying code of the `WebBuilder` function to be extended to incorporate a range of traits that could influence the realisation of potential trophic interactions, other than phylogeny, such as life stage or body size. For instance, within freshwater food webs body size is an important determinant of

trophic interactions, and food web structure predicted using body size alone may be more accurate than those predicted using phylogeny alone (Woodward et al., 2010). This could further increase the realism of the constructed food webs and hence their wider applicability and usefulness.

This dataset of trophic interactions was collated to test the performance of the `WebBuilder` function when predicting the structure of the four empirical UK freshwater food webs used here. It would be straightforward to extend the coverage of this dataset by augmenting it with data collected from other geographic regions. If this dataset is used to construct food webs in the future researchers will need to use their discretion to decide how applicable it is to their system. For instance, this initial version of the dataset does not provide good coverage of lentic species, or species from across Europe or other parts of the world. However, interaction data are being published at a rapidly accelerating rate (Ings et al., 2009) and this can be used to form an iterative feedback process, improving data quality over time; the presence of links predicted by the `WebBuilder` function can therefore be tested evermore rigorously in the future. Identifying underrepresented nodes in the dataset will help target further research more cost effectively: e.g., a great deal is known about the diet of a handful of often economically valuable species in the dataset (for instance, brown trout, *Salmo trutta* appears >3000 times), but very little is known about many others. Additionally, technologies such as those provided by recent advances in molecular sequencing will improve the efficiency of trophic interaction detection (Clare, 2014) and therefore the volume of data which can be incorporated into the dataset. We actively encourage researchers with suitable data to contribute them to this dataset. Exciting initiatives such as Global Biotic Interactions (Poelen et al., 2014), by incorporating necessary information such as the method through which an interaction was determined, could provide a global, open source repository of interaction data which the `WebBuilder` function could access through R. As more of these unknown links become known, nodes will not need to be generalised taxonomically in order to find matches in the dataset, the links generated will more closely match the known links for those species and therefore the quality of the ecological food webs generated by the `WebBuilder` function will improve.

4.3. Conclusions

We have demonstrated that the food webs generated here are comparable to empirically observed food webs and exceeded the accuracy of other potential methods of predicting freshwater food webs. This method could be used to build vast numbers of ecological networks from data that already exists, such as routine biomonitoring data which is collected in huge volumes in many parts of the world (e.g., Dutch soil biomonitoring data have recently been used to build a large collection of food webs; Cohen and Mulder, 2014). Producing collections of replicable networks is vital for advancing ecological network research beyond the largely unreplicated case-study approach that has dominated to date: the `WebBuilder` function approach presents a new robust and repeatable method that helps move us considerably closer to that goal.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.fooweb.2015.09.001.

References

Allesina, S., 2011. predicting trophic relations in ecological networks: a test of the allometric diet breadth model. J. Theor. Biol. 279, 161–168. http://dx.doi.org/10.1016/j.jtbi.2010.06.040.

Allesina, S., Pascual, M., 2009. Food web models: a plea for groups. Ecol. Lett. 12, 652–662. http://dx.doi.org/10.1111/j.1461-0248.2009.01321.x.

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232. http://dx.doi.org/10.1111/j.1365-2664.2006.01214.x.

Barnes, C., 2008. Predator and prey body sizes in marine food webs. Ecology 89, 881.

Bascompte, J., Jordano, P., Melian, C.J., Olesen, J.M., 2003. The nested assembly of plant-animal mutualistic networks. Proc. Natl. Acad. Sci. U. S. A. 100, 9383–9387. http://dx.doi.org/10.1073/pnas.1633576100.

Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., Fox, J.W., Jansen, V.A.A., Iwan Jones, J., Kokkoris, G.D., Logofet, D.O., McKane, A.J., Montoya, J.M., Petchey, O., 2004. Interaction strengths in food webs: issues and opportunities. J. Anim. Ecol. 73, 585–598. http://dx.doi.org/10.1111/j.0021-8790.2004.00833.x.

Briand, F., 1983. Environmental control of food web structure. Ecology 64, 253–263.

Brose, U., Cushing, L., Berlow, E.L., Jonsson, T., Banasek-Richter, C., Bersier, L.-F., ... Martinez, N.D., 2005. Body sizes of consumers and their resources. Ecology 86 (9). http://dx.doi.org/10.1890/05-0379 2545–2545.

Clare, E.L., 2014. Molecular detection of trophic interactions: emerging trends, distinct advantages, significant considerations and conservation applications. Evol. Appl. 7, 1144–1157. http://dx.doi.org/10.1111/eva.12225.

Cleveland, W., Grosse, E., Shyu, W., Chambers, J., Hastie, T., 1992. Local Regression Models. Stat. Model. S.

Cohen, J.E., Mulder, C., 2014. Soil invertebrates, chemistry, weather, human management, and edaphic food webs at 135 sites in The Netherlands: SIZEWEB. Ecology 95, 578.

Cohen, J.E., Newman, C.M., Briand, F., 1985. A stochastic theory of community food webs: II. Individual webs. Proc. R. Soc. B Biol. Sci. http://dx.doi.org/10.1098/rspb.1985.0043.

De Ruiter, P.C., Neutel, A.M., Moore, J.C., 1995. Energetics, patterns of interaction strengths, and stability in real ecosystems. Science 269, 1257–1260. http://dx.doi.org/10.1126/science.269.5228.1257 (80-.).

Edwards, F.K., Lauridsen, R.B., Fernandes, W.P.A., Beaumont, W.R.C., Ibbotson, A.T., Scott, L., Davies, C.E., Jones, J.I., 2009. Re-introduction of atlantic salmon, salmo salar L., to the tadnoll brook, dorset. Proc. Dorset Nat. Hist. Archaeol. Soc. 130, 9–16.

Eklof, a., Helmus, M.R., Moore, M., Allesina, S., 2012. Relevance of evolutionary history for food web structure. Proc. R. Soc. B Biol. Sci. 279, 1588–1596. http://dx.doi.org/10.1098/rspb.2011.2149.

Gilljam, D., Thierry, A., Edwards, F.K., Figueroa, D., Ibbotson, A.T., Jones, J.I., Lauridsen, R.B., Petchey, O.L., Woodward, G., Ebenman, B., 2011. Seeing double: size-based and taxonomic views of food web structure. Adv. Ecol. Res. 45, 67–133. http://dx.doi.org/10.1016/B978-0-12-386475-8.00003-4.

Goldwasser, L., Goldwasser, L., Roughgarden, J., Roughgarden, J., 1993. Construction and analysis of a large caribbean food web. Ecology 74, 1216–1233.

Gravel, D., Poisot, T., Albouy, C., Velez, L., Mouillot, D., 2013. Inferring food web structure from predator–prey body size relationships. Methods Ecol. Evol. http://dx.doi.org/10.1111/2041-210X.12103 (n/a–n/a).

Gray, C., Baird, D.J., Baumgartner, S., Jacob, U., Jenkins, G.B., O'Gorman, E.J., Lu, X., Ma, A., Pocock, M.J.O., Schuwirth, N., Thompson, M., Woodward, G., 2014. Ecological networks: the missing links in biomonitoring science. J. Appl. Ecol. 51, 1444–1449. http://dx.doi.org/10.1111/1365-2664.12300.

Hall, S.J., Raffaelli, D., 1991. Food-web patterns: lessons from a species-rich web. J. Anim. Ecol. 60, 823–841.

Havens, K.E., 1993. Predator–prey relationships in natural community food webs. Oikos 68, 117. http://dx.doi.org/10.2307/3545316.

Henson, K.S.E., Craze, P.G., Memmott, J., 2009. The restoration of parasites, parasitoids, and pathogens to heathland communities. Ecology 90, 1840–1851. http://dx.doi.org/10.1890/07-2108.1.

Hudson, L.N., Emerson, R., Jenkins, G.B., Layer, K., Ledger, M.E., Pichler, D.E., Thompson, M.S.A., O'Gorman, E.J., Woodward, G., Reuman, D.C., 2013. Cheddar: analysis and visualisation of ecological communities in R. Methods Ecol. Evol. 4, 99–104. http://dx.doi.org/10.1111/2041-210X.12005.

Ings, T.C., Montoya, J.M., Bascompte, J., Bluthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E., Lewis, H.M., Olesen, J.M., van Veen, F.J.F., Warren, P.H., Woodward, G., 2009. Ecological networks — beyond food webs. J. Anim. Ecol. 78, 253–269. http://dx.doi.org/10.1111/j.1365-2656.2008.01460.x.

Kaiser-Bunbury, C.N., Muff, S., Memmott, J., Muller, C.B., Caflisch, A., Müller, C.B., 2010. The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour. Ecol. Lett. 13, 442–452. http://dx.doi.org/10.1111/j.1461-0248.2009.01437.x.

Layer, K., Hildrew, A., Monteith, D., Woodward, G., 2010a. Long-Term Variation in the Littoral Food web of an Acidified Mountain Lake. Glob. Chang. Biol. http://dx.doi.org/10.1111/j.1365-2486.2010.02195.x (no–no).

Layer, K., Hildrew, A.G., Woodward, G., 2013. Grazing and detritivory in 20 stream food webs across a broad pH gradient. Oecologia 171, 459–471. http://dx.doi.org/10.1007/s00442-012-2421-x.

Layer, K., Riede, J.O., Hildrew, A.G., Woodward, G., 2010b. Streams Across a Wide pH Gradient Food Web Structure and Stability in 20. Adv. Ecol. Res. 42, 265–299. http://dx.doi.org/10.1016/S0065-2504(10)42005-X.

Ledger, M.E., Brown, L.E., Edwards, F.K., Milner, A.M., Woodward, G., 2012. Drought alters the structure and functioning of complex food webs. Nat. Clim. Chang. 3, 223–227. http://dx.doi.org/10.1038/nclimate1684.

Macfadyen, S., Craze, P.G., Polaszek, A., van Achterberg, K., Memmott, J., 2011. Parasitoid diversity reduces the variability in pest control services across time on farms. Proc. R. Soc. B Biol. Sci. 278, 3387–3394. http://dx.doi.org/10.1098/rspb.2010.2673.

Memmott, J., Martinez, N.D., Cohen, J.E., 2000. Predators, parasitoids and pathogens: species richness, trophic generality and body sizes in a natural food web. J. Anim. Ecol. 69, 1–15. http://dx.doi.org/10.1046/j.1365–2656.2000.00367.x.

Mulder, C., Elser, J.J., 2009. Soil acidity, ecological stoichiometry and allometric scaling in grassland food webs. Glob. Chang. Biol. 15, 2730–2738. http://dx.doi.org/10.1111/j.1365-2486.2009.01899.x.

Olesen, J.M., Bascompte, J., Dupont, Y.L., Jordano, P., 2007. The modularity of pollination networks. Proc. Natl. Acad. Sci. U. S. A. 104, 19891–19896. http://dx.doi.org/10.1073/pnas.0706375104.

Olito, C., Fox, J.W., 2014. Species traits and abundances predict metrics of plant – pollinator network structure, but not pairwise interactions. Oikos 1–9 http://dx.doi.org/10.1111/oik.01439.

Petchey, O.L., Beckerman, A.P., Riede, J.O., Warren, P.H., 2008. Size, foraging, and food web structure. Proc. Natl. Acad. Sci. U. S. A. 105, 4191–4196. http://dx.doi.org/10.1073/pnas.0710672105.

Petchey, O.L., Beckerman, A.P., Riede, J.O., Warren, P.H., 2011. Fit, efficiency, and biology: some thoughts on judging food web models. J. Theor. Biol. 279, 169–171. http://dx.doi.org/10.1016/j.jtbi.2011.03.019.

Piechnik, D.A., Lawler, S.P., Martinez, N.D., 2008. Food-web assembly during a classic biogeographic study: species' "trophic breadth" corresponds to colonization order. Oikos 117, 665–674. http://dx.doi.org/10.1111/j.0030-1299.2008.15915.x.

Pocock, M.J.O., Evans, D.M., Memmott, J., 2012a. The robustness and restoration of a network of ecological networks. Science 335, 973–977. http://dx.doi.org/10.1126/science.1214915 (80-.).

Pocock, M.J.O.O., Memmott, J., Evans, D.M., Memmott, J., 2012b. Supporting Material — the robustness and restoration of a network of ecological networks. Science 335, 973–977. http://dx.doi.org/10.1126/science.1214915 (80-.).

Poelen, J.H., Simons, J.D., Mungall, C.J., 2014. Global biotic interactions: an open infrastructure to share and analyze species-interaction datasets. Ecol. Inform. 24, 148–159. http://dx.doi.org/10.1016/j.ecoinf.2014.08.005.

R Core Team, 2013. R: A Language and Environment for Statistical Computing.

Rall, B.C., Brose, U., Hartvig, M., Kalinkat, G., Schwarzmüller, F., Vucic-Pestic, O., Petchey, O.L., 2012. Universal temperature and body-mass scaling of feeding rates. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 367, 2923–2934. http://dx.doi.org/10.1098/rstb.2012.0242.

Rohr, R.P., Scherer, H., Kehrli, P., Mazza, C., Bersier, L.-F., 2010. Modeling food webs: exploring unexplained structure using latent traits. Am. Nat. 176, 170–177. http://dx.doi.org/10.1086/653667.

Sáyago, R., Lopezaraiza-Mikel, M., Quesada, M., Álvarez-Añorve, M.Y., Cascante-Marín, A., Bastida, J., 2013. Evaluating factors that predict the structure of a commensalistic epiphyte-phorophyte network. Proc. Biol. Sci. 280, 20122821. http://dx.doi.org/10.1098/rspb.2012.2821.

Strong, J.S., Leroux, S.J., 2014. Impact of non-native terrestrial mammals on the structure of the terrestrial mammal food web of newfoundland, Canada. PLoS One 9, e106264. http://dx.doi.org/10.1371/journal.pone.0106264.

Thompson, R.M., Brose, U., Dunne, J.A., Hall, R.O., Hladyz, S., Kitching, R.L., Martinez, N.D., Rantala, H., Romanuk, T.N., Stouffer, D.B., Tylianakis, J.M., 2012. Food webs: reconciling the structure and function of biodiversity. Trends Ecol. Evol. 27, 689–697. http://dx.doi.org/10.1016/j.tree.2012.08.005.

Tylianakis, J.M., Laliberté, E., Nielsen, A., Bascompte, J., 2010. Conservation of species interaction networks. Biol. Conserv. 143, 2270–2279. http://dx.doi.org/10.1016/j.biocon.2009.12.004.

Vázquez, D.P., Chacoff, N.P., Cagnolo, L., 2009. Evaluating multiple determinants of the structure of plant–animal mutualistic networks. Ecology 90, 2039–2046. http://dx.doi.org/10.1890/08-1837.1.

Vazquez, D.P., Melian, C.J., Williams, N.M., Bluthgen, N., Krasnov, B.R., Poulin, R., 2007. Species abundance and asymmetric interaction strength in ecological networks. Oikos 116, 1120–1127. http://dx.doi.org/10.1111/j.2007.0030–1299.15825.x.

Verdú, M., Valiente-Banuet, A., 2011. The relative contribution of abundance and phylogeny to the structure of plant facilitation networks. Oikos 120, 1351–1356. http://dx.doi.org/10.1111/j.1600-0706.2011.19477.x.

Vizentin-Bugoni, J., Maruyama, P.K., Sazima, M., 2014. Processes entangling interactions in communities: forbidden links are more important than abundance in a hummingbird–plant network. Proc. R. Soc. B Biol. Sci. 281. http://dx.doi.org/10.1098/rspb.2013.2397 (20132397–20132397).

Williams, R.J., Martinez, N.D., 2000. Simple rules yield complex food webs. Nature 404, 180–183.

Woodward, G., Blanchard, J., Lauridsen, R.B., Edwards, F.K., Jones, J.I., Figueroa, D., Warren, P.H., Petchey, O.L., 2010. Individual-based food webs: species identity, body size and sampling effects. Adv. Ecol. Res. 43, 211–266. http://dx.doi.org/10.1016/b978-0-12-385005-8.00006-x.

Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman, D.C., Ledger, M.E., 2012. Climate change impacts in multispecies systems: drought alters food web size structure in a field experiment. Philos. Trans. R. Soc. B Biol. Sci. 367, 2990–2997. http://dx.doi.org/10.1098/rstb.2012.0245.

Woodward, G., Speirs, D.C., Hildrew, A.G., 2005. Quantification and resolution of a complex, size-structured food web. Adv. Ecol. Res. 36, 85–135.